

Human Prostate Cancer and Benign Prostatic Hyperplasia: Molecular Dissection by Gene Expression Profiling¹

Jun Luo,² David J. Duggan,² Yidong Chen, Jurga Sauvageot, Charles M. Ewing, Michael L. Bittner, Jeffrey M. Trent, and William B. Isaacs³

Brady Urological Institute, Johns Hopkins Medical Institutions, Baltimore, Maryland 21287-2101 [J. L., J. S., C. M. E., W. B. I.], and Cancer Genetics Branch, National Human Genome Research Institute, Bethesda, Maryland 20892-4470 [D. J. D., Y. C., M. L. B., J. M. T.]

Abstract

Critical aspects of the biology and molecular basis for prostate malignancy remain poorly understood. To reveal fundamental differences between benign and malignant growth of prostate cells, we performed gene expression profiling of primary human prostate cancer and benign prostatic hyperplasia (BPH) using cDNA microarrays consisting of 6500 human genes. Frozen prostate specimens were processed to facilitate extraction of RNA from regions of tissue enriched in either benign or malignant epithelial cell growth within a given specimen. Gene expression in each of the 16 prostate cancer and nine BPH specimens was compared with a common reference to generate normalized measures for each gene across all of the samples. Using an analysis of complete pairwise comparisons of expression profiles among all of the samples, we observed clearly discernable patterns of overall gene expression that differentiated prostate cancer from BPH. Further analysis of the data identified 210 genes with statistically significant differences in expression between prostate cancer and BPH. These genes include many not recognized previously as differentially expressed in prostate cancer and BPH, including *hepsin*, which codes for a transmembrane serine protease. This study reveals for the first time that significant and widespread differences in gene expression patterns exist between benign and malignant growth of the prostate gland. Gene expression analysis of prostate tissues should help to disclose the molecular mechanisms underlying prostate malignant growth and identify molecular markers for diagnostic, prognostic, and therapeutic use.

Introduction

Adenocarcinoma of the prostate gland is the most common form of malignancy diagnosed in the United States male, accounting for over 35% of all of the cancers affecting men (1). Approximately 20% of those diagnosed will eventually die from this disease. Prostate cancer progression is a process involving multiple molecular alterations (2, 3), many of which can be reflected in changes of gene expression in the prostate carcinoma cells. BPH,⁴ on the other hand, is the most common benign tumor in men >60 years of age (4). Benign growth of the prostate gland is accompanied by a significant increase in the proliferation rate of epithelial cells in the hyperplastic acini (5). Because these epithelial cells actively proliferate but do not frequently progress to malignancy, they serve as a very useful cell population for comparison with prostate carcinoma cells (6). Therefore, comparative analysis of gene expression in prostate cancer and BPH specimens may provide important information relating to malignant transformation of prostatic cells. Additionally, a systematic gene expression

analysis of this kind is very likely to facilitate the identification of molecular markers and therapeutic targets for the improved management of prostate cancer patients.

The emerging technology of cDNA microarrays provides the ability to comparatively analyze mRNA expression of thousands of genes in parallel (7, 8). Previous studies (9–11) have revealed novel features of human cancers by classifying tumors based on gene expression profiles. Human gene expression patterns derived from cDNA microarray measurements have been increasingly used to identify genes associated with human malignancies in a number of organ sites (12–16). On the basis of these studies, it seems apparent that cDNA microarray-based gene expression analysis of human prostate tissues, especially those from well-documented clinical sources, would reveal molecular characteristics associated with prostate tumorigenesis. In this study, we obtained gene expression profiles of 16 primary prostate cancers and nine BPH specimens. A complete pairwise comparison of the 25 samples revealed consistently distinctive patterns of gene expression between these two groups of prostate tissues. Statistical tools were used to identify genes with sufficient discriminative power to differentiate these two groups of samples, generating a list of genes with significantly different expression levels between malignant growth and benign growth of the prostate gland.

Materials and Methods

Prostate Tissue Specimens. Prostate cancer tissue specimens were obtained from 16 patients undergoing radical prostatectomy for clinically localized prostate carcinoma at Johns Hopkins Hospital from October 1998 and March 2000. Seven of the nine BPH specimens were obtained from patients undergoing open prostatectomy and two from patients undergoing transurethral resection of the prostate at Johns Hopkins Hospital between February 1999 and November 2000. Harvested tissues were flash frozen in liquid nitrogen and stored at -80°C until use. Specimens were chosen for analysis according to two criteria: (a) sufficient tissue was available for analysis; and (b) histological evaluation by H&E staining demonstrated that the samples contained predominantly epithelial cells (in the case of BPH samples) or adenocarcinoma cells (in the case of the cancer samples; see below). Frozen tissue blocks were trimmed after histological evaluation to meet this latter criterion. Institutional Review Board-approved informed consent was obtained from all of the patients in this study.

RNA Preparation. Trimmed prostate blocks were cut into 10- μm sections in a cryostat. A total of 200 frozen sections/specimen were cut and maintained on dry ice for RNA extraction. Sectioning of the samples facilitates subsequent tissue homogenization, ensuring the maximum quality and yield of RNA preparations. In addition, the first and last sections from each specimen were preserved for pathological confirmation and calculation of percentages of tumor and epithelium. Efforts were made to enrich the epithelial composition in the samples. After trimming, the 16 tumor specimens contained at least 60% (range, 60–85%) adenocarcinoma cells in cellular composition. Six of the seven BPH samples from open prostatectomy contain at least 50% (range, 50–70%) epithelial cells, whereas the two BPH samples obtained by transurethral resection were 40% and 45% in epithelial content. Detailed tissue data are

Received 2/28/01; accepted 5/1/01.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¹ Supported by Public Health Service SPORE CA 58236 and DK 52675.

² These authors participated equally.

³ To whom requests for reprints should be addressed, at 115 Marburg, 600 N. Wolfe Street, Johns Hopkins Hospital, Baltimore, MD 21287. Phone: (410) 955-2518; Fax: (410) 955-0833; E-mail: wisaacs@jhmi.edu.

⁴ The abbreviations used are: BPH, benign prostatic hyperplasia; MDS, multidimensional scaling; RT-PCR, reverse transcription-PCR.

provided in supplemental information.⁵ Total RNA was isolated as described (9). Briefly, the aqueous portion from the Trizol/chloroform (Life Technologies, Inc., Rockville, MD) extraction step was mixed with equal volume of 70% ethanol and loaded on a Qiagen Rneasy (Qiagen, Valencia, CA) column. The columns were then processed according to manufacturer's recommendations. RNA samples were subsequently concentrated using Microcon 100 concentrators (Millipore, Bedford, MA) to the desired concentration and stored at -80°C until use.

Array Fabrication. The 6500 sequence-verified human cDNAs, representing 6112 unique genes (4573 known genes) on the basis of Unigene build 128, were obtained under a Cooperative Research and Development Agreement with Research Genetics. A complete annotated list of these cDNAs is available from the supplemental information.⁵ Printing of the cDNA clones was carried out as described previously (9). Briefly, amplified fragments from the clones were printed onto poly-L-lysine-coated glass slides. One week after printing, the arrayed slides were UV radiated to cross-link the DNA targets, treated with succinic anhydride to block poly-L-lysine, and boiled to denature DNA targets.

Fluorescent Labeling and Hybridization. Labeling of total RNA was achieved by direct incorporation of Cy5-dUTP or Cy3-dUTP (Amersham Pharmacia, Piscataway, NJ) in a reverse transcription reaction using anchored oligodeoxythymidylate primer (Genosys, The Woodlands, TX) and Superscript II reverse transcriptase (Life Technologies, Inc.). Fluor-tagged cDNAs were then concentrated to the desired volume using Microcon concentrators (Millipore). Detailed labeling procedures are available from the website.⁶ For each of the 25 surgical samples, Cy3-dUTP-tagged cDNAs were mixed with Cy5-dUTP-tagged common reference (Fig. 1) and subsequently cohybridized to a microarray. A single reference sample composed of a pool of RNA from two BPH specimens was used throughout all of the hybridizations to ensure normalized measures for each gene in each individual sample.

Image Analysis and Data Collection. Hybridized slides were scanned using the Axon GenePix4000A scanner (Axon Instruments, Foster City, CA), and images were processed using a collection of IPLab (Scanalytics, Inc., Fairfax, VA) extensions developed at the Cancer Genetics Branch at National Human Genome Research Institute (17). The image processing analysis now also extracts information regarding spot quality and assigns a quality score to each ratio measurement, with 0 as the lowest measurement quality and 1 as the highest measurement quality. The definition of the quality metric is based on the notion that unreliable data points usually result from weak target intensity, high local background, small target area, and inconsistent target intensity within a given target. Implementation of the quality metric enables unified and universally applicable data filtering before downstream higher-level data analysis. Meanwhile, computation of the similarity measures can be easily modified by introducing the quality score into the calculation without prior data filtering as shown below. Details of the quality metric are provided in supplemental information.⁵

Data Analysis. The similarity between gene expression patterns is measured by computing the Euclidean distances for each pair of samples based on log-transformed ratios across all of the genes (18). Calculation of the Euclidean distance between sample x and y , d_{xy} , was modified by introducing the quality score into the equation to yield

$$d_{xy} = \sqrt{\frac{\sum_{i=1}^n w_i (x_i - y_i)^2}{\sum_{i=1}^n w_i}}$$

where x_i and y_i represent the log-transformed expression ratio of i th gene in sample x and y , respectively (total of n genes in each sample), and $w_i = q_{xi}q_{yi}$, where q_{xi} and q_{yi} are the expression measurement quality for i th gene in the sample x and y , respectively. Using a matrix of Euclidean distance measurements from the complete pairwise comparison of all of the prostate specimens, a multidimensional scaling method (MDS) (9, 19) was used to display the overall similarity in gene expression profiles. During the MDS procedure, samples were positioned in a three-dimensional space so that the distance between each pair of samples very closely approximates the Euclidean distance measurements in the matrix for the corresponding sample pair. This three-dimensional approximation of multidimensional relationships produces a visually intuitive pattern of sample clustering. Weighted gene analysis was

performed to yield a list of genes statistically significant in separating BPH and prostate tumor (9, 20). Briefly, for two groups (prostate cancer and BPH) with a given number of samples 16 and 9, the discriminative weight for each gene $w = d_B / (k_1 d_{w1} + k_2 d_{w2} + \alpha)$; where d_B is the between group Euclidean distance, d_{w1} is the average Euclidean distance among all of the prostate samples, d_{w2} is the average Euclidean distance among all of the BPH samples, $k_1 = 16/(16 + 9)$, $k_2 = 9/(16 + 9)$, and α is a small constant to ensure the denominator is never equal to zero. Genes are ranked according to the w value. Genes with high w values create greater separation between groups and denser compaction within the groups; i.e., they have more discriminative power to differentiate the two groups. To test the statistical significance of the discriminative weights, sample labels were randomly permuted (9, 20) among the two groups, and the w value for each gene was again computed. This random permutation of sample labels was repeated 1000 times to generate a w distribution that would be expected under the assumption of random gene expression; i.e., no difference between the groups. The w values generated from the actual data were then assigned P s based on the w distribution of randomized data. An agglomerative hierarchical clustering algorithm (9) based on Euclidean distance measure was used to cluster the genes with statistically significant ($P < 0.001$) w values; i.e., genes statistically different in expression between prostate cancer and BPH samples.

RT-PCR. Expression of the *hepsin* gene was verified using RT-PCR in six prostate cancer samples and six BPH samples randomly chosen from the 25 prostate tissue specimens. The cDNA synthesis was performed following the manufacturer's instructions (Roche Molecular Biochemical, Indianapolis, IN) using a primer set for *hepsin* (forward, gatgtctgcaatggcgtgac; reverse, ccacacgccgcgaacgtg). Prostate-specific antigen (forward, ccacaccgctctacga; reverse, ttgatccactccggaatgc) was used as a control for equal amount of prostate epithelial cells represented in each loading.

Results

A total of 25 frozen prostate tissue specimens (16 prostate cancer and nine BPH samples) collected at the time of surgery were analyzed in this study. A quality control measure was applied to ensure that the samples were enriched in epithelial content by trimming, sectioning, and subsequent histological review within each specimen. Total RNA was extracted, and fluorescently labeled cDNA probes (Cy3-labeled) prepared from each of these samples were cohybridized to the arrayed targets along with the common reference probe (Cy5-labeled) derived from a pool of two BPH specimens (Fig. 1). Normalized fluorescent intensity ratios from each hybridization experiment represent the relative mRNA abundance for each gene in each sample compared with the common reference. Analysis of the extent of similarity of the gene expression ratios between samples then provided a measure of the overall similarity in gene expression patterns between samples. A complete pairwise comparison of all of the samples was performed by computing the Euclidean distance for each pair of samples based on all of the log-transformed ratios. The quality score associated with each ratio measurement was incorporated into the calculation to ensure that the Euclidean distance measurements were not sensitive to unreliable data points with low quality score, which is typically a result of low signal intensity value and small target size. A matrix of Euclidean distances from a complete pairwise comparison was generated. To create a visual representation of relationships among all of the samples in terms of their similarities in gene expression profiles, a three-dimensional mapping of the samples, where the Euclidean distance between samples was closely approximated by the inter-sample map distances, was created using a multidimensional scaling method (9, 19). Samples that have gene expression profiles that are more similar to each other will lie closer and form aggregation (cluster) in three-dimensional space. As seen in this plot (Fig. 2A), a strong distinction in the pattern of overall gene expression is evident between prostate tumor samples (blue) and BPH (golden brown) samples (see supplemental information⁵ for a three-dimensional animation of the MDS plot). Samples within each group showed similar

⁵ http://www.nhgri.nih.gov/DIR/Microarray/Prostate_Supplement.

⁶ <http://www.nhgri.nih.gov/DIR/microarray>.

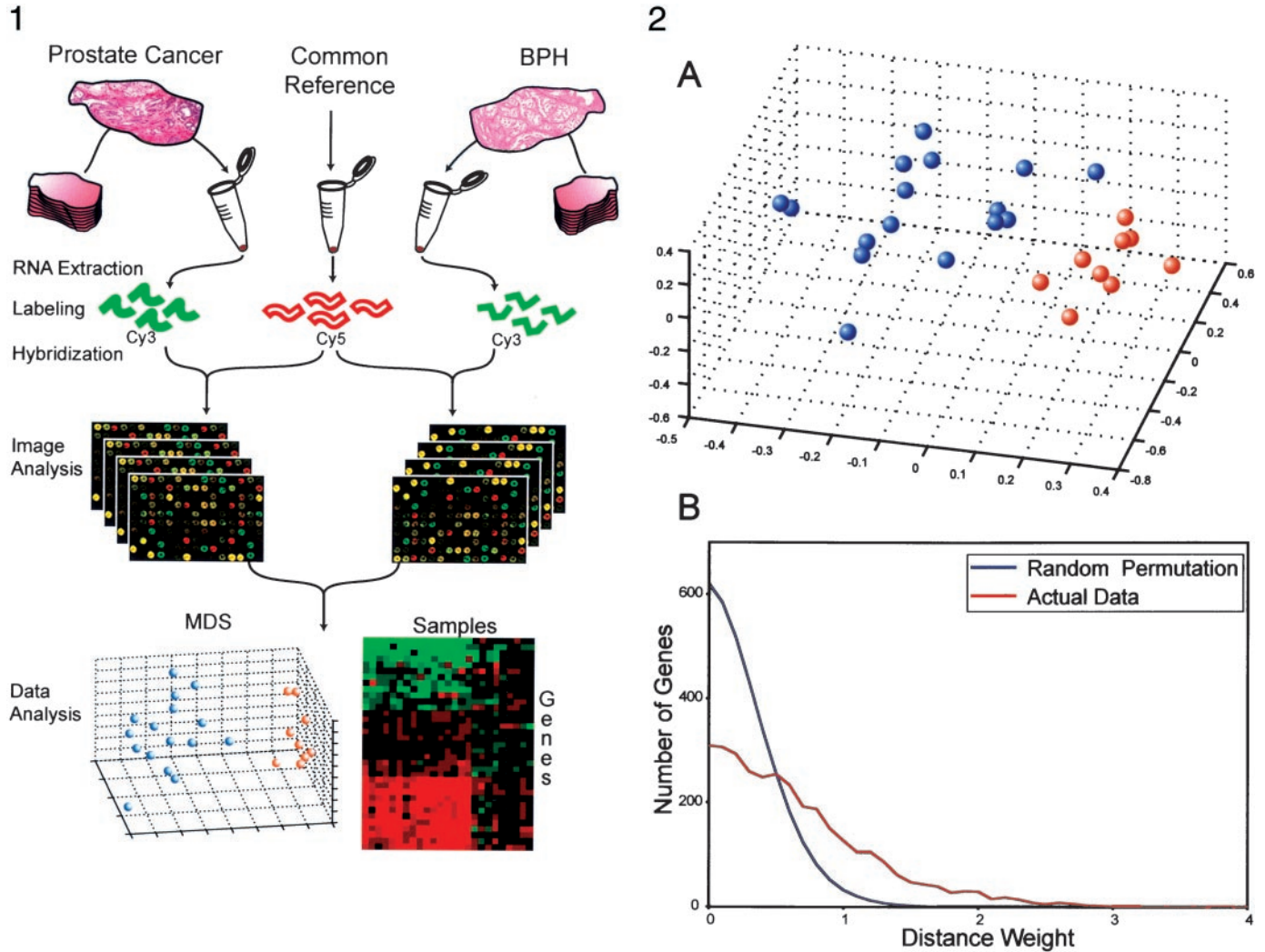


Fig. 1. Overview of experimental procedures for gene expression profiling of prostate tissues. Prostate samples were trimmed and sectioned to enrich epithelial content in each specimen and to facilitate sample homogenization. Total RNA was extracted from the samples and labeled with Cy3-dUTP in a RT reaction. RNA from a pool of two BPH specimens was labeled in parallel with Cy5-dUTP and used as reference sample for all of the 16 prostate cancer and nine BPH samples. Labeled products from the test samples were mixed with the labeled reference and cohybridized to microarrays containing cDNAs for 6500 human genes. Images were scanned, and data were analyzed to study the gene expression patterns.

Fig. 2. MDS and weighted gene analysis. A, a three-dimensional plot produced by a MDS method displaying the overall similarity in gene expression among the 25 samples. The 16 prostate cancer samples are shown in *blue*, and the nine BPH samples are shown in *golden brown*. B, distribution of the w values from weighted gene analysis. The w values are on the X axis, and the number of genes are on the Y axis. For the actual data (*red line*), the w value for each corresponding gene represents the degree of difference in expression of that gene between two defined groups, 16 prostate cancer samples *versus* nine BPH samples. The *blue line* represents the w distribution that would be expected under the null hypothesis, which assumes no difference in gene expression between the two groups. C, a three-dimensional plot of the 25 prostate samples based on the top 210 genes from weighted gene analysis. The 16 prostate tumor samples are in *blue*, the nine BPH samples are in *golden brown*.

gene expression patterns by forming a localized grouping of BPH samples that is readily separable from the cancer sample grouping. This result indicates that it is possible to draw a distinction between benign growth and malignant growth of prostatic cells solely based on the overall similarity of gene expression patterns.

To determine which gene expression patterns exhibited the greatest difference between BPH and prostate cancer samples, weighted gene

analysis (9, 20) was performed. This analysis generates an ordered list of genes with statistically significant differences in expression between BPH and prostate cancer. Filtering out unreliable ratio measurements results in a set of 3215 genes for weighted gene analysis. First, the w value for each gene was computed to analyze the discriminative power of that gene to separate prostate cancer and BPH; *i.e.*, the difference in expression of that gene between prostate cancer and

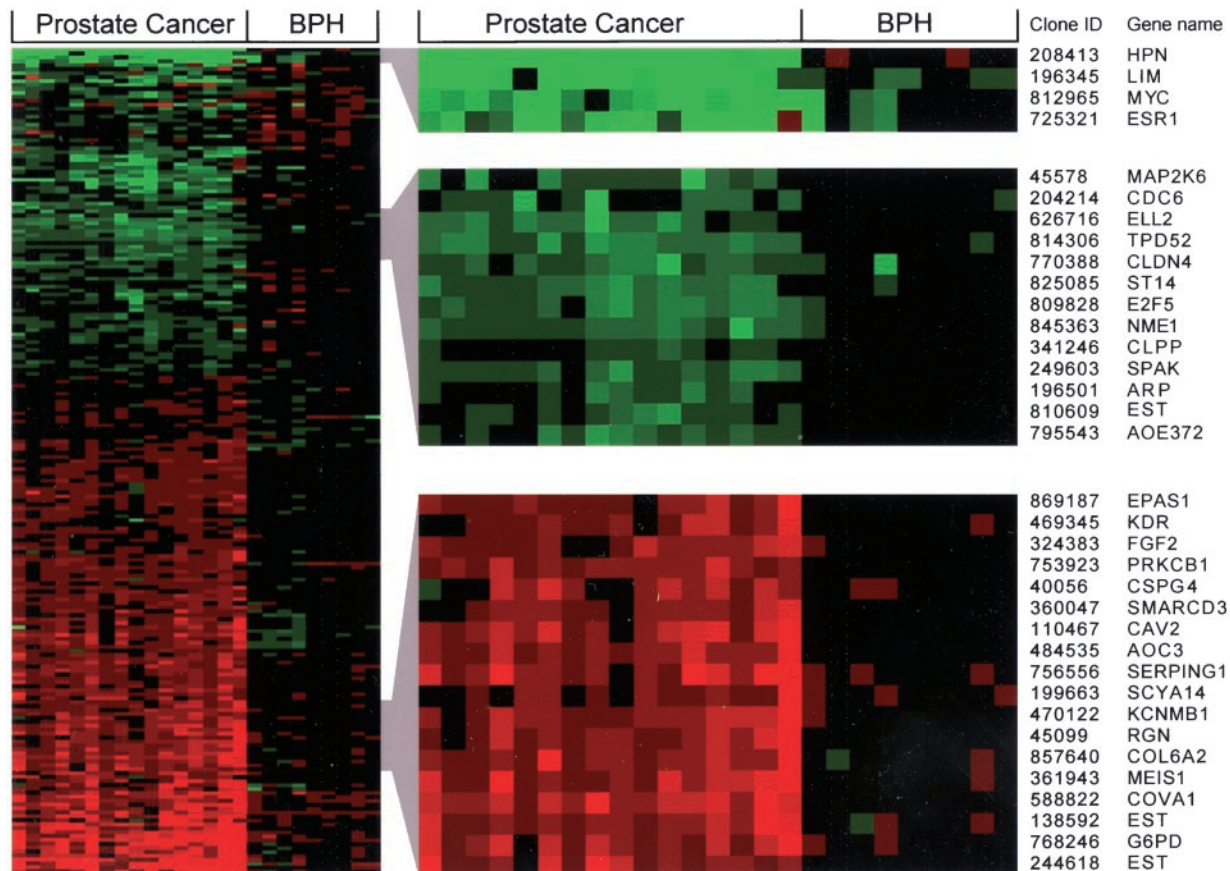


Fig. 3. The top 210 genes with a statistically significant difference in expression between prostate cancer and BPH. Data are organized in a matrix format following hierarchical clustering analysis of the 210 genes. Each *row* represents a single gene; each *column* represents a prostate sample. Normalized ratios correlating to the abundance of mRNA relative to a common reference are represented by colors; *red*, down-regulated relative to reference; *green*, up-regulated relative to reference; *black*, approximately same as reference. Color saturation represents the magnitude of deviation from the reference. Selected clusters of genes are listed with corresponding gene symbol and IMAGE clone ID.

BPH. Genes were then ranked according to w values, with the largest w value indicating the most discriminative power to separate prostate cancer from BPH. A fitted line representing the w distribution from the actual data is displayed in Fig. 2B (*red line*). Next, a w distribution was created from the randomly permuted gene expression data sets (Fig. 2B, *blue line*), representing the w distribution that would be expected under the null hypothesis that no true difference exists between the two groups. Therefore, each w value from the actual data can be assigned a P to determine the statistical significance of the associated gene to differentiate prostate cancer from BPH, by corresponding the w value (from the actual data) to the w distribution from the randomized data. Genes with w value above a critical value 1.7 were determined to be statistically significant ($P < 0.001$) in expression between prostate cancer and BPH (see supplemental information⁵ for details). As shown in Fig. 2B, it is apparent that the observed gene expression difference between prostate cancer and BPH is not the result of random events. There are 210 genes with w values > 1.7 (and thus $P < 0.001$) from the actual dataset (*red line*), whereas no gene in the random datasets has a w value > 1.7 (*blue line*). An MDS plot was created to visualize the relationships among the 25 samples based on these 210 genes (Fig. 2C). As expected, a greater degree of separation was observed because this list of genes represents the subset of genes with the most expression differences between BPH and prostate cancer samples.

The 210 genes are clustered and displayed in Fig. 3 along with their relative expression in each sample compared with a common reference. Samples are ordered as groups of prostate cancer and BPH to facilitate visual comparison of the expression levels. The measured

expression ratios for each gene are presented graphically as colored images, with the *green squares* (rectangular in compressed image) representing higher expression in sample compared with the reference, the *red squares* meaning lower expression in sample than reference, and the *black squares* indicating a ratio of approximately 1. Color intensities are scaled according to the ratio (reference:sample), with the brightest color having a ratio of greater than 5 (*red*) or smaller than 0.2 (*green*). For clarity of data presentation, we only list three clusters of genes with their associated names and IMAGE clone ID numbers (Fig. 3). A complete list of the 210 genes with associated clustering tree and other details can be accessed from supplemental information.⁵

The 210 genes with w values > 1.7 can be ranked according to w values. The number one ranked gene (*i.e.*, having the greatest ability to differentiate BPH from cancer) is *hepsin* ($w = 5.05$), which codes for a transmembrane serine protease that has been implicated in cell growth, development, and initiation of blood coagulation, and is overexpressed in ovarian cancer (21). This gene was found to be highly expressed in prostate cancer samples relative to BPH samples (Fig. 3; first gene). RT-PCR analysis was used to determine the expression level of the *hepsin* gene in six prostate cancer samples and six BPH samples. Prostate-specific antigen, a prostate luminal epithelial marker and also a serine protease, was used as a loading control as well as an indicator of the epithelial content in the samples. We confirmed the high expression of *hepsin* in prostate tumor samples, whereas minimal or no signal was detected in BPH samples (Fig. 4).

Many of the differentially expressed genes remain to be confirmed independently. However, some of them can be indirectly verified by

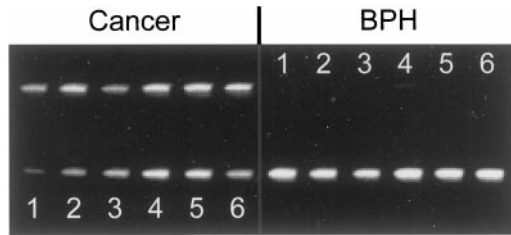


Fig. 4. Verification of *hepsin* expression in BPH and prostate cancer by RT-PCR. Top row, *hepsin*; bottom row, prostate-specific antigen.

searching the public National Center for Biotechnology Information serial analysis of gene expression database.⁷ For example, database searching of the 34 genes (excluding *hepsin*) from the three gene clusters shown in Fig. 3 returned 17 genes with available data on relative expression in prostate cancer *versus* that in normal prostate. Strikingly, 12 of the 17 genes were confirmed to be differentially expressed with reasonable confidence (at least 2-fold change in serial analysis of gene expression data derived from PR317 prostate libraries).⁷ This observation also suggests that gene expression changes between prostate cancer and BPH reflect in large degree the differences between normal and cancerous prostate epithelium. Additional efforts will be needed to fully characterize the expression levels of these 210 genes in benign and malignant prostate tissues.

Discussion

This study was undertaken as a step toward discovering some of the fundamental differences between benign and malignant growth of prostate epithelial cells. The comparison of BPH and prostate cancer is thought likely to lead toward a more incisive understanding of the biology of tumors because BPH appears to occupy a state that is unusually close to that of prostate cancer; both involve overgrowth of the epithelial cells. Whereas cancerous growth of the prostate epithelial cells is characterized by accumulation of molecular abnormalities because of genomic instability, BPH represents overgrowth of a more "normal epithelium" with rare genetic abnormalities (22). Thus, it is expected that many of the differences that can be observed between BPH and cancerous epithelia will reflect this particular aspect of prostate tumor biology. The tool chosen to carry out the comparison was gene expression profiling using cDNA microarrays. Mathematical analysis of the profiling results demonstrated that clear differences in expression pattern can be seen both at the overall expression level (Fig. 2A) and at the individual gene level (Fig. 3).

Interpretation of the observed differences is bound both by the complex nature of cellular heterogeneity and by our knowledge of the tissue origin for BPH and prostate cancer. Any comparison is limited by the homogeneity of the samples being compared. A typical surgical prostate tissue specimen usually presents a mixture of different cell types, each with a potentially unique gene expression profile. The prostate samples used in this study were processed to maximize the percentage of the target epithelia from which RNA was extracted to reduce the contributions of the contaminating tissues to the final profiles. The likelihood of the observed differences in expression representing differences in BPH and prostate cancer biology is further heightened by using multiple samples. The contaminating tissues will be more randomly represented in the samples analyzed, and their contribution to the analysis will thus be further diluted. The other source of expression differences between BPH and prostate tumor samples that may be tangential to the cancer-specific differences is the tissue of origin of the two sample types. BPH and prostate cancer are

pathological entities arising in two different areas of the prostate gland (22). The majority (~80%) of prostate cancers are found in the peripheral zone, and almost all of the BPH occurs in a periurethral region, termed the transition zone. Clarification of the expression differences that arise from the differences in normal peripheral and transition zone tissue will require studies of the relative expression of genes of interest in these tissues. Nevertheless, many genes that are consistently up-regulated and down-regulated in the majority of prostate cancer samples when compared with BPH are most likely representative of molecular features associated with prostate malignancy. On the other hand, future studies focusing on the identification of genes that have expression that is zone-specific should shed light on the mechanisms underlying the regional difference in the incidence of benign and malignant growth of the prostatic cells.

Genomic instability of prostate tumors could lead to an extensive variation in gene expression even within a single tumor (2). Therefore, identification of tumor-specific gene expression changes common to all of the tumors is of particular interest; *e.g.*, mRNA expression of the *hepsin* gene is strikingly high in all of the prostate cancer samples compared with minimal expression in all of the BPH samples examined. Although it is not clear at this point what implications this gene as well as the other highly discriminating genes might have on prostate malignancies, the cellular function of the gene products and the potential use of those malignancy-associated genes as molecular markers warrants further study.

Although important features of prostate tumor biology remain to be investigated by including additional prostate tissue samples differing in pathological characteristics, the current study reports both a clear overall and gene-by-gene difference between gene expression profiles associated with malignant growth and benign growth of the prostatic cells. This study is currently being expanded by using microarrays containing more genes known to be important in prostate biology and by reanalysis of the profiles as new sample sets are added. Analysis of the roles of the genes already suggested as possibly important in prostate cancer and the further development of profiles of the various types of normal and cancerous prostate epithelia offer a reasonable approach to developing an understanding of the biology of prostate malignancy.

Acknowledgments

We thank Arthur Glatfelter, Chris Gooden, and Spyro Mousses for microarray technical assistance. We also thank Dr. Angelo De Marzo for valuable suggestions regarding the manuscript and Darryl Leja for help with scientific illustration.

References

1. Ries, L. A. G., Wingo, P. A., Miller, D. S., Howe, H. L., Weir, H. K., Rosenberg, H. M., Vernon, S. W., Cronin, K., and Edwards, B. K. The annual report to the nation on the status of cancer, 1973–1997, with a special section on colorectal cancer. *Cancer* (Phila.), 88: 2398–2424, 2000.
2. Isaacs, W. B., and Bova, G. S. Prostate cancer. In: B. Vogelstein and K. Kinzler (eds.), *The Genetic Basis of Human Cancer*, pp. 653–660. New York: McGraw-Hill, 1998.
3. Thompson, T. C., Timme, T. L., Bangma, C. H., Nasu, Y., Hull, G. W., Hall, S. J., and Stapleton, A. M. F. Molecular biology of prostate cancer. In: N. J. Vogelzang, P. T. Scardino, W. U. Shipley, and D. S. Coffey (eds.), *Genitourinary Oncology* (Ed. 2), pp. 563–564. Philadelphia: LWW, 2000.
4. Ziada, A., Rosenblum, M., and Crawford, E. D. Benign prostatic hyperplasia: an overview. *Urology*, 53 (Suppl. 3A): 1–6, 1999.
5. Kyprianou, N., Tu, H., and Jacobs, S. C. Apoptotic *versus* proliferative activities in human benign prostatic hyperplasia. *Hum. Pathol.*, 27: 668–675, 1996.
6. Cole, K. A., Krizman, D. B., and Emmert-Burk, M. R. The genetics of cancer—a 3D model. *Nat. Genet.*, 21 (Suppl. 1): 38–41, 1999.
7. Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* (Wash. DC), 270: 467–470, 1995.
8. Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P., and Trent, J. M. Expression profiling using cDNA microarrays. *Nat. Genet.*, 21 (Suppl. 1): 10–14, 1999.

⁷ <http://www.ncbi.nlm.nih.gov/SAGE/SAGEcid.cgi?>

9. Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., Sampas, N., Dougherty, E., Wang, E., Marincola, F., Gooden, C., Lueders, J., Glatfelter, A., Pollock, P., Carpten, J., Gillanders, E., Leja, D., Dietrich, K., Beaudry, C., Berens, M., Alberts, D., Sondak, V., Hayward, N., and Trent, J. Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature (Lond.)*, **406**: 536–540, 2000.
10. Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. C., Sabet, H., Tran, T., Yu, X., Powell, J. I., Yang, L., Marti, G. E., Moore, T., Hudson, J., Jr., Lu, L., Lewis, D. B., Tibshirani, R., Sherlock, G., Chan, W. C., Greiner, T. C., Weisenburger, D. D., Armitage, J. O., Warnke, R., Levy, R., Wilson, W., Grever, M. R., Byrd, J., Botstein, D., Brown, P. O., and Staudt, L. M. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature (Lond.)*, **403**: 503–511, 2000.
11. Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A. L., Brown, P. O., and Botstein, D. Molecular portraits of human breast tumors. *Nature (Lond.)*, **406**: 747–752, 2000.
12. Ono, K., Tanaka, T., Tsunoda, T., Kitahara, O., Kihara, C., Okamoto, A., Ochiai, K., Takagi, T., and Nakamura, Y. Identification by cDNA microarray of genes involved in ovarian carcinogenesis. *Cancer Res.*, **60**: 5007–5011, 2000.
13. Sgroi, D. C., Teng, S., Robinson, G., LeVangie, R., Hudson, J. R., Jr., and Elkahoul, A. G. *In vivo* gene expression profile analysis of human breast cancer progression. *Cancer Res.*, **59**: 5656–5661, 1999.
14. Sallinen, S. L., Sallinen, P. K., Haapasalo, H. K., Helin, H. J., Helen, P. T., Schraml, P., Kallioniemi, O. P., and Kononen, J. Identification of differentially expressed genes in human gliomas by DNA microarray and tissue chip techniques. *Cancer Res.*, **60**: 6617–6622, 2000.
15. Shim, C., Zhang, W., Rhee, C. H., and Lee, J. H. Profiling of differentially expressed genes in human primary cervical cancer by complementary DNA expression array. *Clin. Cancer Res.*, **4**: 3045–3050, 1998.
16. Xu, J., Stolk, J. A., Zhang, X., Silva, S. J., Houghton, R. L., Matsumura, M., Vedvick, T. S., Leslie, K. B., Badaro, R., and Reed, S. G. Identification of differentially expressed genes in human prostate cancer using subtraction and microarray. *Cancer Res.*, **60**: 1677–1682, 2000.
17. Chen, Y., Dougherty, E. R., and Bittner, M. L. Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J. Biomed. Opt.*, **2**: 364–374, 1997.
18. Johnson, R. A., and Wichern, D. W. *Applied Multivariate Statistical Analysis* (Fourth Ed.), pp. 28. New Jersey: Prentice Hall, 1998.
19. Khan, J., Simon, R., Bittner, M., Chen, Y., Leighton, S. B., Pohida, T., Smith, P. D., Jiang, Y., Gooden, G. C., Trent, J. M., and Meltzer, P. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. *Cancer Res.*, **58**: 5009–5013, 1998.
20. Hedenfalk, I., Duggan, D., Chen, Y., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O., Wilfond, B., Borg A., and Trent, J. Gene-expression profiles in hereditary breast cancer. *N. Engl. J. Med.*, **344**: 539–548, 2001.
21. Tanimoto, H., Yan, Y., Clarke, J., Korourian, S., Shigemasa, K., Parmley, T. H., Parham, G. P., and O'Brian, T. J. Hepsin, a cell surface serine protease identified in hepatoma cells, is overexpressed in ovarian cancer. *Cancer Res.*, **57**: 2884–2887, 1997.
22. De Marzo, A., Coffey, D. S., and Nelson, W. G. New concepts in tissue specificity for prostate cancer and benign prostatic hyperplasia. *Urology*, **53** (Suppl. 3A): 29–40, 1999.